An Interpretable Data-Driven Approach for Decision-Making in Materials Discovery Amidst Uncertainties

ID: ORL02
Oral Talk
Aug. 23, 14.20 - 14.40

Minh-Quyet Ha and Hieu-Chi Dam
Japan Advanced Institute of Science and Technology, Nomi, Japan
Email: dam@jaist.ac.jp

Keywords: Material informatics, Decision-making, Uncertainty modelling, Dempster-Shafer theory

Advanced materials exhibit exceptional properties and complex structures due to their multiple constituent elements, offering a wide range of potential candidates for new material development. However, the experimental validation and simulation of these candidates demand significant resources. Therefore, material scientists must make informed decisions about candidate selection while managing resource constraints. The primary challenge lies in evaluating the rationality of each decision, balancing two key objectives: exploitation and exploration. Exploitation focuses on maximizing immediate benefits by selecting candidates with high-performance potential based on existing knowledge. Exploration, on the other hand, seeks to expand understanding by choosing lesser-known or underexplored candidates. To quantitatively assess the rationality of these decisions, it is essential to model the uncertainties associated with their outcomes. Two main types of uncertainties are considered: aleatoric uncertainty, which arises from inherent randomness and unpredictable individual events, and epistemic uncertainty, which results from a lack of knowledge or data and can be reduced with additional information. Exploitation typically favors candidates with a higher probability of success but must address aleatoric uncertainty. Exploration broadens the understanding of the materials space, contending with aleatoric and epistemic uncertainties. This research introduces a data-driven approach that models the uncertainties of decision outcomes, supporting scientists in making rational choices. The framework employs quantitative methods to assess material similarity, providing a basis for analogy-based, intuitive, and interpretable reasoning. This approach is further supported by a quantitative model for uncertainty grounded in the Dempster-Shafer theory, a generalization of Bayesian theory, enabling the quantitative measurement of both aleatoric and epistemic uncertainties.